

**UNITED STATES DISTRICT COURT
SOUTHERN DISTRICT OF NEW YORK**

THE NEW YORK TIMES COMPANY,

Plaintiff,

v.

MICROSOFT CORPORATION, OPENAI,
INC., OPENAI LP, OPENAI GP, LLC,
OPENAI, LLC, OPENAI OPCO LLC,
OPENAI GLOBAL LLC, OAI
CORPORATION, LLC, OPENAI
HOLDINGS, LLC,

Defendants.

Civil Action No. 1:23-cv-11195-SHS

Hon. Sidney H. Stein

**REPLY MEMORANDUM OF LAW IN
SUPPORT OF OPENAI'S OBJECTION
TO NONDISPOSITIVE ORDER
PURSUANT TO FEDERAL RULE OF
CIVIL PROCEDURE 72(a)**

I. INTRODUCTION

The Times is engaged in a transparent effort to curtail its discovery productions to ensure that it produces only those materials that support its claims. But discovery “is not a one-way street.” *Lee v. Delta Airlines, Inc.*, 2023 WL 9184679, at *9 (E.D.N.Y. Mar. 14, 2023). OpenAI has sought discovery that is critical to its fair use defense and necessary to rebut the Times’s claims of existential harm to its ability to provide “[i]ndependent journalism.” ECF 170 (“FAC”) ¶¶ 1, 2, 5. The Times has exclusive possession of the revenue, traffic, and licensing information that OpenAI seeks, and such materials are essential for OpenAI to establish that the Times has not suffered any cognizable market harm for purposes of the fair use defense. The requested discovery thus goes to the very heart of the critical issues in this case, and to one of OpenAI’s primary defenses. While Supreme Court and Second Circuit case law compel the production of such materials, the orders at issue (ECF 352 & 353, the “December Orders”) failed to engage with that case law and relied instead on a prior discovery order addressing entirely different issues. This Court’s intervention is necessary to ensure that such a critical case is decided upon a complete, and fair, record.

II. ARGUMENT

A. The Times manufactures *ex post facto* justifications for the December Orders.

As explained in OpenAI’s opening papers, Magistrate Judge Wang denied both of OpenAI’s market harm motions without directly responding to the parties’ arguments. ECF 372 (“Mem.”) 8–9. Indeed, the December Orders were summary orders denying the requests for pre-motion conferences “for the reasons stated in ECF 344.” But the discovery order at ECF 344 did not directly address the legal bases for the requested materials. On the contrary, it considered unrelated discovery requests regarding the Times’s development and/or use of non-OpenAI

generative AI tools. *See* Mem. 9 (describing the November 22 order).¹ Recognizing this problem, the Times manufactures justifications for what *may* have motivated the Magistrate Judge’s orders. The Times has no response to the procedural reality—it cannot claim after the fact that the Magistrate Judge based her decision on something other than what her orders actually said. Regardless, as described more fully below, the Times’s *ex post facto* justifications are not persuasive.

B. The Revenue and Traffic Requests seek documents critical to OpenAI’s fair use defense.

1. The Times misstates the scope of the Factor Four inquiry.

As OpenAI explained, documents and communications regarding (1) alleged decreases in the Times’s revenue, including the reasons for those decreases (OpenAI RFPs 141, 146–48); and (2) alleged decreases in subscriptions and traffic to the Times’s website and mobile applications, including the reasons for those decreases (OpenAI RFPs 35, 37, 142–45) (together, the “Revenue and Traffic Requests”) are evidence critical to OpenAI’s ability to show that its products are not “the source of the loss[es]” alleged by the Times under factor four of the fair use defense. Mem. 11–15 (quoting *Google LLC v. Oracle Am., Inc.*, 593 U.S. 1, 35 (2021)).² If the Times’s purported market loss has nothing to do with OpenAI—but instead pandemic disruptions, social media, cable news, or podcasts—then that evidence weighs strongly in favor of fair use. This is not only common sense; it is the law. *See, e.g., Harper & Row Publishers, Inc. v. Nation Enter.*,

¹ OpenAI has separately objected to the November 22, 2024 order. *See* ECF 362–363.

² The Times also failed to respond to OpenAI’s argument that the Revenue and Traffic Requests are relevant to damages for similar reasons. Mem. 11–15 (citing, *e.g., Fitzgerald Publ’g Co. v. Baylor Publ’g Co.*, 807 F.2d 1110, 1118 (2d Cir. 1986) (“[T]he primary measure of recovery is the extent to which the market value of the copyrighted work at the time of the infringement has been injured or destroyed *by the infringement*.” (emphasis added))). This is an additional, independent basis for ordering the requested discovery.

471 U.S. 539, 567 (1985) (defendant has the burden “to show that [any] damage would have occurred had there been no taking of copyrighted expression”); *Hachette Book Grp., Inc. v. Internet Archive*, 115 F.4th 163, 191 (2d Cir. 2024) (“[T]he defendant may present evidence of a lack of market harm to support their defense of fair use.”).

The Times’s opposition fails to substantively engage with these common-sense propositions. Instead, it cherry-picks citations from these cases to argue—wrongly—that the fourth factor fair use analysis depends on looking at only a *defendant’s* use in a vacuum. *See* ECF 432 (“Opp.”) 10–11 (characterizing the cases as holding that the fourth factor looks “at the infringing use itself” and that evidence of whether a substitute was created would “come primarily from [d]efendants”). The cases say no such thing; in fact, they say the opposite. Fair use cases clearly state that evidence from both parties is relevant. *See, e.g., Oracle*, 593 U.S. at 36, 40 (evidence that plaintiff Oracle was “poorly positioned to succeed in the mobile phone market” and that “Sun was beset by business challenges in developing a mobile phone product” both “weigh[ed] in favor of fair use”); *see also Hachette*, 115 F.4th at 191–92 (faulting defendant for relying on third-party data instead of plaintiff’s data, where the magistrate judge had rejected defendant’s motion to compel such data from plaintiff). While the *burden* to establish a fair use defense lies with OpenAI, *see id.*, it is the Times that possesses relevant *evidence* regarding trends in subscriptions, revenue, and traffic that Open AI needs given its burden. To claim, as the Times does, that assessing whether a plaintiff has suffered “harm to the actual or potential markets for the copyrighted work” requires only scrutiny of the *defendant’s* conduct is absurd on its face. *See* Opp. at 10 (citing *Oracle*, 593 U.S. at 24). The question is whether *the Times* has suffered harm, and the Times itself uniquely possesses such evidence.

2. The Times’s selective productions are not proportional to the needs of this case.

Recognizing the frailty of its legal arguments, the Times pivots to arguing that it has allegedly already agreed to provide the relevant discovery, going so far as to append a long list of document categories it has already agreed to produce in response to other requests. *See* Opp. 3–4, 11 & Ex. 1. The Times’s claim that it has already agreed to produce responsive materials is a concession that the requested evidence is relevant. But the Times’s productions to date are manifestly inadequate: even a cursory review of the categories the Times cites reveals that the Times has agreed only to produce documents that it believes will support its claims, not those that are relevant to OpenAI’s fair use defense. For example, the Times has agreed to produce what it says in its opposition are “documents concerning how *Defendants* have diverted Times readers and online traffic” and “documents regarding decreases in online traffic or impacts on Wirecutter subscriptions and advertising revenue *attributed to Defendants*.” Opp. 3–4 (emphases added). But these commitments parrot the Times’s allegations in its complaint. *See* FAC ¶¶ 129 (alleging that OpenAI products “[d]ecreased traffic” to Wirecutter and thus “le[d] to a loss of revenue for Wirecutter”), 157 (alleging that OpenAI’s products “divert readers . . . away from” the Times’s websites). The Times has categorically refused to produce any documents or communications relating to *other* potential or actual causes of alleged revenue or subscription loss. With its refusal, the Times is essentially asking this Court to endorse a presumption that any alleged loss is necessarily traceable to OpenAI. That cannot stand.

3. The Times cannot avoid producing market harm-related discovery by redefining its theory of harm.

The Times next contends that its promised production of causation analyses regarding reduced website traffic obviates the need to produce materials concerning any other alleged market harm or the causes thereof. *See* Opp. 11. But the Times’s allegations of harm are not so

limited. The Times’s complaint alleges losses not just to website traffic, but also to advertising revenues, subscriptions, and affiliate link revenues. FAC ¶¶ 154–157. Website traffic is simply one narrow category among multiple that are relevant to the Times’s multiple alleged harms. Moreover, while website traffic may (or may not) be associated with revenue, it is not direct evidence of market harm. On the contrary, the critical question is whether such alleged decreases in traffic have resulted in *economic* harm, which is precisely what the Times is refusing to produce. The Times apparently intends to argue that a decrease in traffic alone presumptively establishes economic harm, while refusing to produce the direct evidence that would support or undermine that claim. Indeed, the Times’s own reporting reveals that such a presumption is improper. The Times recently reported “crossing the threshold of 11 million total subscribers for the first time” and a 16.1% increase in profit in the third quarter of 2024, compared to the same period in 2023, calling into question its allegations that it has suffered a loss of traffic, revenue, and subscriptions due to OpenAI’s products.³

Indeed, even if traffic to the Times’s website has decreased, the reasons behind the decrease and whether the decrease resulted in market harm are critical evidence to the fair use analysis. The Times’s cabined commitment to produce traffic analyses to the exclusion of all others has forced the Times to take absurd positions. According to the Times, if a Times analysis concluded that all of the Times’s alleged revenue fluctuations were caused by post-pandemic cultural shifts, the Times’s position is that such an analysis need not be produced in this litigation. But that very evidence would lend overwhelming support to OpenAI’s fair use defense because it would establish that the harm “would have occurred had there been no taking

³ Katie Robertson, *New York Times Passes 11 Million Subscribers*, N.Y. Times (Nov. 4, 2024), <https://www.nytimes.com/2024/11/04/business/media/new-york-times-earnings.html>.

of copyrighted expression.” *Harper & Row*, 471 U.S. at 567. The market harm inquiry cannot be reduced to an inquiry into internet traffic to a single platform; the Times’s refusal to produce a full set of documents responsive to these requests remains unjustified.

4. The Times has failed to specifically articulate, and the Court did not consider, burden.

Finally, the Times’s characterizations of OpenAI’s requests as “expansive demand[s] for every last document or email regarding changes in traffic to the Times’s website or its revenue for any reason” ring hollow. *See* Opp. 11. First, the Times still has not substantiated the purported burden it faces. *See Zhulinska v. Niyazov Law Grp., P.C.*, 2021 WL 5281115, at *3 (E.D.N.Y. Nov. 12, 2021) (party cannot evade discovery obligations by offering “merely speculative” concerns about burden). In any event, OpenAI has never demanded production of “every last document” related to revenue or traffic, nor would it do so. To avoid producing relevant documents, the Times is pretending that standard e-discovery practices, like the application of search terms and custodial limitations, do not exist. As with any e-discovery search, these processes and limitations will allow OpenAI access to relevant evidence—such as emails or communications in custodial repositories discussing reasons for trends in Times revenue, traffic, or subscriptions—in a manner proportional to the Times’s allegations in this case. OpenAI is prepared to meet and confer on such strategies, but the Times has refused to do so. The Times’s unsubstantiated burden claims are not grounds for denying this discovery, and it should be ordered to produce documents in response to the Revenue and Traffic Requests.

C. The Licensing Requests seek documents critical to OpenAI’s fair use defense.

The Times alleges that (1) it has sustained harm to a “well-established market” for licensing its content for commercial purposes that generates millions of dollars in revenue for the Times per year, FAC ¶¶ 5, 155–157; and (2) that there is a potential licensing market for using

the Times’s works for training LLMs, *id.* ¶ 92. The December Orders erroneously denied OpenAI’s discovery requests for documents and communications relating to (i) the Times’s efforts to license the asserted works (RFPs 18, 33); (ii) the market for licensing or using the Times’s works for AI training, including documents showing or refuting whether such a market exists (RFP 160); and (iii) whether the Times’s works have ever been used for the purpose of analyzing language patterns (RFP 161). Similar to the Revenue and Traffic Requests, these documents are relevant because they allow OpenAI to challenge the Times’s allegations of harm to its licensing market, both actual and potential. Mem. 15–18 (citing, *e.g.*, *Bill Graham Archives v. Dorling Kindersley, Ltd.*, 448 F.3d 605, 614 (2d Cir. 2006)). The Times’s responses miss the mark. Opp. 14–15.

1. The Times refuses to produce critical discovery regarding the alleged existence of a licensing market for its Asserted Works as training data.

The Times has refused to produce documents and communications related to whether a licensing market for LLM training exists, cabining its production to documents related to specific licensing efforts. Opp. at 15. Yet the Times simultaneously seeks production of precisely the same documents it refuses to produce itself. On January 13, 2025, the Times filed a letter motion seeking to compel Microsoft to produce “documents that address the existence of a market for [training] data, *regardless of whether such documents address a particular negotiation with a particular partner.*” ECF 396 at 2 (emphasis added). It further argued that such documents “go to the heart of the fourth factor of the fair use analysis,” and that “[t]here is simply no excuse for withholding these plainly relevant documents.” *Id.* And yet that is precisely what the Times is doing here, endeavoring to create a one-sided fair use record by parroting its own allegations. The Times has undeniably conceded both the relevance and proportionality of such materials, so it has no basis to resist producing them.

In addition, the Times fails to respond to evidence demonstrating that its efforts to license its works for training LLMs are strategic, litigation-driven efforts as opposed to reasonable, commercially motivated efforts to enter a new legitimate market (which would be relevant to the market harm inquiry). Mem. 17–18. As OpenAI explained, “courts have recognized limits on the concept of ‘potential licensing revenues’ by considering only traditional, reasonable, or likely to be developed markets”; the absence of a “ready market or means to pay for the use” weighs in favor of fair use. Mem. 5 (quoting *Am. Geophysical Union v. Texaco, Inc.*, 60 F.3d 913, 929, 931 (2d Cir. 1994)). The Times cites no authority to the contrary. This evidence is necessary to allow the Court to fully evaluate the Times’s narrative that a purported market for licensing Times works for LLMs was a legitimate market, rather than a manufactured one in service of the Times’s allegations in this lawsuit.

2. The Times refuses to produce critical discovery related to its allegations of harm to its preexisting licensing markets.

As for discovery into the Times’s existing licensing of its works, the Times misrepresents what OpenAI seeks and why. Contrary to its protests that “[t]here is simply no way” that these documents are relevant, Opp. 15, it was the Times that put its broader licensing scheme at issue by alleging that OpenAI’s products affect its *preexisting* licensing revenues. FAC ¶¶ 5, 157. The Times points to its limited commitments to produce the license agreements themselves and some communications regarding attempts to license the Times’s works to Defendants in this case, but as with the other requests at issue here, those documents provide no contextual discovery into whether OpenAI’s products have any connection to the alleged losses. And here again the Times ignores standard e-discovery techniques to alleviate purported burden. As with the Revenue and Traffic Requests, the Times never bothered to engage in the search-term or

custodian-identification process, so its complaints now ring hollow. OpenAI has been ready to engage in that process but has been consistently rebuffed.

The Times should be ordered to produce documents in response to the Licensing Requests.

III. CONCLUSION

For the foregoing reasons, OpenAI respectfully requests that this Court vacate the December Orders and order the Times to produce the requested discovery.

KEKER, VAN NEST & PETERS LLP

Dated: January 27, 2025

By: /s/ Andrew Dawson

ROBERT A. VAN NEST (*pro hac vice*)
R. JAMES SLAUGHTER (*pro hac vice*)
PAVEN MALHOTRA - # 4409397
MICHELLE S. YBARRA (*pro hac vice*)
NICHOLAS S. GOLDBERG (*pro hac vice*)
THOMAS E. GORMAN (*pro hac vice*)
KATIE LYNN JOYCE (*pro hac vice*)
ANDREW DAWSON (*pro hac vice*)
SARAH SALOMON (*pro hac vice*)
633 Battery Street
San Francisco, CA 94111-1809
Telephone: 415 391 5400
Facsimile: 415 397 7188
rvannest@keker.com
rslaughter@keker.com
pmalhotra@keker.com
mybarra@keker.com
ngoldberg@keker.com
tgorman@keker.com
kjoyce@keker.com
adawson@keker.com
ssalomon@keker.com

Attorneys for Defendants OPENAI, INC.,
OPENAI LP, OPENAI GP, LLC, OPENAI,
LLC, OPENAI OPCO LLC, OPENAI
GLOBAL LLC, OAI CORPORATION,
LLC, OPENAI HOLDINGS, LLC

Dated: January 27, 2025

LATHAM & WATKINS LLP

By: /s/ Elana Nightingale Dawson*

Andrew M. Gass (*pro hac vice*)

andrew.gass@lw.com

Joseph R. Wetzel

joseph.wetzel@lw.com

505 Montgomery Street, Suite 2000

San Francisco, CA 94111

Telephone: 415.391.0600

Sarang V. Damle

sy.damle@lw.com

Elana Nightingale Dawson (*pro hac vice*)

elana.nightingaledawson@law.com

Michael David (*pro hac vice*)

michael.david@law.com

555 Eleventh Street, NW, Suite 1000

Washington, D.C. 20004

Telephone: 202.637.2200

Allison L. Stillman

alli.stillman@lw.com

Rachel R. Blitzer

rachel.blitzer@lw.com

Herman Yue

herman.yue@lw.com

Luke A. Budiardjo

luke.budiardjo@lw.com

1271 Avenue of the Americas

New York, NY 10020

Telephone: 212.906.1200

Attorneys for OpenAI Defendants

Dated: January 27, 2025

MORRISON & FOERSTER LLP

By: /s/ Rose S. Lee*

Joseph C. Gratz (*pro hac vice*)
jgratz@mofo.com
425 Market Street
San Francisco, CA 94105-2482
Telephone: 415.268.7000

Rose S. Lee (*pro hac vice*)
rose.lee@mofo.com
707 Wilshire Boulevard, Suite 6000
Los Angeles, California 90017-3543
Telephone: 213.892.5200

Eric K. Nikolaidis
enikolaides@mofo.com
Emily C. Wood
ewood@mofo.com
250 West 55th Street
New York, NY 10019-9601
Telephone: 212.468.8000

Carolyn Homer (*pro hac vice*)
chomer@mofo.com
2100 L Street, NW Suite 900
Washington, D.C. 20038
Telephone: 202.887.1500

Attorneys for OpenAI Defendants

* All parties whose electronic signatures are included herein have consented to the filing of this document in accordance with Rule 8.5(b) of the Court's ECF Rules and Instructions.

CERTIFICATE OF WORD COUNT COMPLIANCE

I certify that the foregoing brief complies with the requirements stated in Local Civil Rule 7.1, and that the brief contains 2,701 words, excluding the parts of the brief exempted from the word count by Local Civil Rule 7.1.

Dated: January 27, 2025

/s/ Andrew Dawson
Andrew Dawson